## SEQUENCE LISTING

<110> MERKULOV, Gennady et al. <120> ISOLATED HUMAN PROTEASE PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND USES THEREOF <130> CL001177DIV2 <140> (to be assigned) <141> 2001-12-14 <150> 09/813,819 <151> 2001-03-22 <150> 09/920,048 <151> 2001-08-02 <160> 4 <170> FastSEQ for Windows Version 4.0 <210> 1 <211> 2968 <212> DNA <213> Homo sapiens <400> 1 cgacctggcc gccggccgct cctccgcgcg ctgttccgca cttgctgccc tcgcccggcc 60 eggagegeeg etgecatgeg getggegetg etetgggeee tggggeteet gggegeggge 120 agecetetge ettectggee geteceaaat atagecetge tgtegattee etcagtactg 180 tettggggtg teetgggace tgeaggtgge actgaggage ageaggeaga gteagagaag 240 gccccgaggg agcccttgga gccccaggtc cttcaggacg atctcccaat tagcctcaaa 300 aaggtgette agaccagtet geetgageee etgaggatea agttggaget ggaeggtgae 360 agtcatatcc tggagctgct acagaatagg gagttggtcc caggccgccc aaccctggtg 420 tggtaccagc ccgatggcac tcgggtggtc agtgagggac acactttgga gaactgctgc 480 taccagggaa gagtgcgggg atatgcaggc tcctgggtgt ccatctgcac ctgctctggg 540 ctcagaggct tggtggtcct gaccccagag agaagctata ccctggagca ggggcctggg 600 gaccttcagg gtcctcccat tatttcgcga atccaagatc tccacctgcc aggccacacc 660 tgtgccctga gctggcggga atctgtacac actcagacgc caccagagca ccccctggga 720 cagcgccaca ttcgccggag gcgggatgtg gtaacagaga ccaagactgt ggagttggtg 780 attgtggctg atcactcgga ggcccagaaa taccgggact tccagcacct gctaaaccgc 840 acactggaag tggccctctt gctggacaca ttcttccggc ccctgaatgt acgagtggca 900 ctagtgggcc tggaggcctg gacccagcgt gacctggtgg agatcagccc aaacccagct 960 gtcaccctcg aaaacttcct ccactggcgc agggcacatt tgctgcctcg attgccccat 1020 gacagtgccc agctggtgac tggtacttca ttctctgggc ctacggtggg catggccatt 1080 cagaactcca tetgttetee tgaettetea ggaggtgtga acatggaeca etecaecage 1140 atcctqqqaq tcqcctcctc cataqcccat qaqttqqqcc acaqcctqqq cctqqaccat 1200 gatttgcctg ggaatagctg cccctgtcca ggtccagccc cagccaagac ctgcatcatg 1260 gaggeeteea cagactteet accaggeetg aactteagea actgeageeg acgggeeetg 1320 gagaaagccc tcctggatgg aatgggcagc tgcctcttcg aacggctgcc tagcctaccc 1380 cctatggctg ctttctgcgg aaatatgttt gtggagccgg gcgagcagtg tgactgtggc 1440 ttcctggatg actgcgtcga tccctgctgt gattctttga cctgccagct gaggccaggt 1500 gcacagtgtg catctgacgg accetgttgt caaaattgcc agetgcgccc gtetggctgg 1560 cagtgtcgtc ctaccagagg ggattgtgac ttgcctgaat tctgcccagg agacagctcc 1620

cagtgtcccc ctgatgtcag cctaggggat ggcgagccct gcgctggcgg gcaagctgtg 1680

tgcatgcacg ggcgttgtgc ctcctatgcc cagcagtgcc agtcactttg gggacctgga 1740 gcccagcccg ctgcgccact ttgcctccag acagctaata ctcggggaaa tgcttttggg 1800 agctgtgggc gcaaccccag tggcagttat gtgtcctgca cccctagaga tgccatttgt 1860 gggcagctcc agtgccagac aggtaggacc cagcctctgc tgggctccat ccgggatcta 1920 ctctgggaga caatagatgt gaatgggact gagctgaact gcagctgggt gcacctggac 1980 ctgggcagtg atgtggccca gccctcctg actctgcctg gcacagcctg tggccctggc 2040 ctggtgtgta tagaccatcg atgccagcgt gtggatctcc tgggggcaca ggaatgtcga 2100 agcaaatgcc atggacatgg ggtctgtgac agcaacaggc actgctactg tgaggagggc 2160 tgggcacccc ctgactgcac cactcagctc aaagcaacca gctccctgac cacagggctg 2220 ctcctcagcc tcctggtctt attggtcctg gtgatgcttg gtgccagcta ctggtaccqt 2280 qcccqcctqc accaqcqact ctqccaqctc aaqqqaccca cctqccaqta caqqqcaqcc 2340 caatctggtc cctctgaacg gccaggacct ccgcagaggg ccctgctggc acgaggcact 2400 aaggetagtg ctctcagett cccqqccccc ccttccagqc cgctqccqcc tgaccctgtg 2460 tecaagagae tecagtetea ggggeeagee aageeeceae eeceaaggaa geeaetgeet 2520 geogaecece agggeoggtg eccategggt gacetgeecg geocagggge tggaateceg 2580 cccctagtgg taccctccag accagcgcca ccgcctccga cagtgtcctc gctctacctc 2640 tgacctetee ggaggtteeg etgeeteeaa geeggaetta gggetteaag aggegggegt 2700 gccctctgga gtcccctacc atgactgaag gcgccagaga ctggcggtgt cttaagactc 2760 egggeacege caegegetgt caageaacae tetgeggace tgeeggegta gttgeagegg 2820 gggcttgggg aggggctggg ggttggacgg gattgaggaa ggtccgcaca gcctgtctct 2880 aaaaaaaaa aaaaaaaaa aaaaaaaa

<210> 2 <211> 855 <212> PRT <213> Homo sapiens

<400> 2

Pro Leu Pro Ser Trp Pro Leu Pro Asn Ile Ala Leu Leu Ser Ile Pro Ser Val Leu Ser Trp Gly Val Leu Gly Pro Ala Gly Gly Thr Glu Glu Gln Gln Ala Glu Ser Glu Lys Ala Pro Arg Glu Pro Leu Glu Pro Gln 55 Val Leu Gln Asp Asp Leu Pro Ile Ser Leu Lys Lys Val Leu Gln Thr 70 75 Ser Leu Pro Glu Pro Leu Arg Ile Lys Leu Glu Leu Asp Gly Asp Ser 90 His Ile Leu Glu Leu Leu Gln Asn Arg Glu Leu Val Pro Gly Arg Pro 100 105 110 Thr Leu Val Trp Tyr Gln Pro Asp Gly Thr Arg Val Val Ser Glu Gly 120 125 His Thr Leu Glu Asn Cys Cys Tyr Gln Gly Arg Val Arg Gly Tyr Ala 135 Gly Ser Trp Val Ser Ile Cys Thr Cys Ser Gly Leu Arg Gly Leu Val 150 155 Val Leu Thr Pro Glu Arg Ser Tyr Thr Leu Glu Gln Gly Pro Gly Asp 170 175 Leu Gln Gly Pro Pro Ile Ile Ser Arg Ile Gln Asp Leu His Leu Pro 180 185 190 Gly His Thr Cys Ala Leu Ser Trp Arg Glu Ser Val His Thr Gln Thr 195 200 Pro Pro Glu His Pro Leu Gly Gln Arg His Ile Arg Arg Arg Asp 220 210 215

Met Arg Leu Ala Leu Leu Trp Ala Leu Gly Leu Leu Gly Ala Gly Ser

Val Val Thr Glu Thr Lys Thr Val Glu Leu Val Ile Val Ala Asp His 230 235 Ser Glu Ala Gln Lys Tyr Arg Asp Phe Gln His Leu Leu Asn Arg Thr 245 250 Leu Glu Val Ala Leu Leu Asp Thr Phe Phe Arg Pro Leu Asn Val 265 Arg Val Ala Leu Val Gly Leu Glu Ala Trp Thr Gln Arg Asp Leu Val Glu Ile Ser Pro Asn Pro Ala Val Thr Leu Glu Asn Phe Leu His Trp 295 300 Arg Arg Ala His Leu Leu Pro Arg Leu Pro His Asp Ser Ala Gln Leu 310 315 Val Thr Gly Thr Ser Phe Ser Gly Pro Thr Val Gly Met Ala Ile Gln 325 330 Asn Ser Ile Cys Ser Pro Asp Phe Ser Gly Gly Val Asn Met Asp His 340 345 Ser Thr Ser Ile Leu Gly Val Ala Ser Ser Ile Ala His Glu Leu Gly 360 His Ser Leu Gly Leu Asp His Asp Leu Pro Gly Asn Ser Cys Pro Cys 375 380 Pro Gly Pro Ala Pro Ala Lys Thr Cys Ile Met Glu Ala Ser Thr Asp 390 Phe Leu Pro Gly Leu Asn Phe Ser Asn Cys Ser Arg Arg Ala Leu Glu 405 410 Lys Ala Leu Leu Asp Gly Met Gly Ser Cys Leu Phe Glu Arg Leu Pro 425 Ser Leu Pro Pro Met Ala Ala Phe Cys Gly Asn Met Phe Val Glu Pro 440 445 Gly Glu Gln Cys Asp Cys Gly Phe Leu Asp Asp Cys Val Asp Pro Cys 455 460 Cys Asp Ser Leu Thr Cys Gln Leu Arg Pro Gly Ala Gln Cys Ala Ser 470 475 Asp Gly Pro Cys Cys Gln Asn Cys Gln Leu Arg Pro Ser Gly Trp Gln 485 490 Cys Arg Pro Thr Arg Gly Asp Cys Asp Leu Pro Glu Phe Cys Pro Gly 505 Asp Ser Ser Gln Cys Pro Pro Asp Val Ser Leu Gly Asp Gly Glu Pro 520 Cys Ala Gly Gly Gln Ala Val Cys Met His Gly Arg Cys Ala Ser Tyr 535 540 Ala Gln Gln Cys Gln Ser Leu Trp Gly Pro Gly Ala Gln Pro Ala Ala 550 555 Pro Leu Cys Leu Gln Thr Ala Asn Thr Arg Gly Asn Ala Phe Gly Ser 565 570 Cys Gly Arg Asn Pro Ser Gly Ser Tyr Val Ser Cys Thr Pro Arg Asp 580 585 Ala Ile Cys Gly Gln Leu Gln Cys Gln Thr Gly Arg Thr Gln Pro Leu 600 Leu Gly Ser Ile Arg Asp Leu Leu Trp Glu Thr Ile Asp Val Asn Gly 615 Thr Glu Leu Asn Cys Ser Trp Val His Leu Asp Leu Gly Ser Asp Val 630 635 Ala Gln Pro Leu Leu Thr Leu Pro Gly Thr Ala Cys Gly Pro Gly Leu 650 Val Cys Ile Asp His Arg Cys Gln Arg Val Asp Leu Leu Gly Ala Gln 660 665

```
Glu Cys Arg Ser Lys Cys His Gly His Gly Val Cys Asp Ser Asn Arg
                            680
His Cys Tyr Cys Glu Glu Gly Trp Ala Pro Pro Asp Cys Thr Thr Gln
    690
                        695
                                            700
Leu Lys Ala Thr Ser Ser Leu Thr Thr Gly Leu Leu Ser Leu Leu
                    710
                                        715
Val Leu Val Leu Val Met Leu Gly Ala Ser Tyr Trp Tyr Arg Ala
                725
                                    730
Arg Leu His Gln Arg Leu Cys Gln Leu Lys Gly Pro Thr Cys Gln Tyr
                                745
            740
                                                    750
Arg Ala Ala Gln Ser Gly Pro Ser Glu Arg Pro Gly Pro Pro Gln Arg
                            760
Ala Leu Leu Ala Arg Gly Thr Lys Ala Ser Ala Leu Ser Phe Pro Ala
                        775
                                            780
Pro Pro Ser Arg Pro Leu Pro Pro Asp Pro Val Ser Lys Arg Leu Gln
                    790
                                        795
Ser Gln Gly Pro Ala Lys Pro Pro Pro Pro Arg Lys Pro Leu Pro Ala
                                    810
                805
Asp Pro Gln Gly Arg Cys Pro Ser Gly Asp Leu Pro Gly Pro Gly Ala
                                825
Gly Ile Pro Pro Leu Val Val Pro Ser Arg Pro Ala Pro Pro Pro
                            840
                                                845
Thr Val Ser Ser Leu Tyr Leu
    850
                        855
<210> 3
<211> 17138
<212> DNA
```

```
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<222> (1)...(17138)
<223> n = A,T,C or G
```

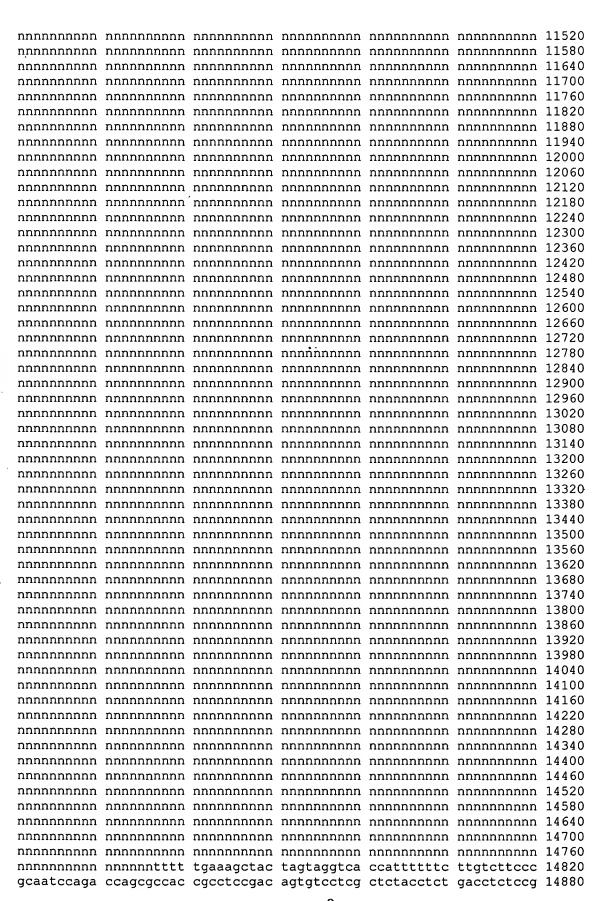
<400> 3

```
ttgggtgacc ctgggcagtg atcacatctc caagcatcag ttttctcacc tgaaaaaaag 60
gagatgataa taacactatc tgccttacat qacaattgaa ttgaattttt tttttttt 120
tgagactaag teteactetg tegeceagge tggagtgeag tggegtgate ttggeteact 180
gcaacctcca cctccccagt tcaagcgatt ctcgtgcctc agcttcccga gtagctggga 240
ttacaggcac acactaccac gcccggctaa tttagaattg aaataattta tgtacagtat 300
cttagtacag gacctgacat tataaacaat gagtggcagc cattcttatt taatcagtcc 360
taacaaagtt cataaaagtg agactgtgtt tgcttagctt tttccctagg gcctggatac 420
ccccagccc catgacacac aataggggcc aaatgaatgt gttgtgaaaa aatgaaaaac 480
aaaaaacaaa aaagaacatg ctgggattcc ttgacagggt cgtgaagcaa actgaatgtg 540
aatgcacaga tggaaatgtg ccagacagtc attccaagca gaatgtgcaa agactcagtc 600
cacagggaat gcgaagtgcc agggctagtc tcaggagaaa ctggctcaga agagacagct 660
ctcagggagg gctaaagtag gaaagaggct aqaaaqqqac caggtgaggg aaggctctga 720
aggccaagcc caagagttct gcctgtctqq caqqcaqcaq ggcctctqqa gtttcttgqq 780
caaagagtgg ctgcttcctg ggtaaggtgg cctgtggaaa atccctgaca actgtgtaga 840
gacatgtcgt gagggatggc agggagcata gtgaactagg tttgtggttt ggaatcaggg 900
cccctggggt ccagccaagt tggattgttt actatctgtg tgactttgag agtcacttca 960
cetttetcaa etgtaaagtg gggatagcaa cagtgatagt egatetggee tgeteaette 1020
tcagcctcac tgtgagaacc aaataagatg atttacagga aagtgcaaat gagagttgtg 1080
getgatatee gettggagag ageetggagg gtgcateete ceatteteea teacagagtt 1140
ggggagggag gcaccetege ectecagggg ttteetttgt ccaacceage etectecaae 1200
```

acgcgggaat tgtcaggcct ggcgacttca gacaggaaac gctgtccagt tccccttctt 1260 tecegeeteg eteceggget ggegetaaeg eccacetece aacagegeea eccgetggeg 1320 gatatectge accgeggetg eccgeteetg egeegetgge tgtgeeggeg etgegtggtg 1380 tgccaggcac ccgagacgcc cgagtcctac gtgtgccgga cgctggactg cgaggccgtg 1440 tactgctggt cgtgctggga cgacatgcgg cagcggtgcc cggtctgcac gccccgcgaa 1500 gagetetett ceteegeett tagtgacage aacgaegaca etgeetaege ggggtgaaga 1560 ggcgtcctgc tcgctcttcc gcaccgtcct tcccggttaa taaaatgccc tgtacgcttc 1620 acgtgggtcg gggactgggg tgagccgcgc actqcctcqc ctgcagtcgg gaaagcctqc 1680 ccgcccgacc tctccgagcc aggccgcgca caggaggcag ggaggccgcg aagctactag 1740 ggaggggtcc ggacctggcg ccgggtgaag gcgcgccgcc caagccggtc ggaccgggca 1800 ccggctccca ctccgcacag ttgcggggaa gcggtagcgc tgagcagcgc gggcgtagtg 1860 ggcggtgtcc ccgctcccga ggcacccggc gcgcagcggg gcgggctttg ccgggggcgg 1920 agettggett ggggcegggt gggaggggc gggcegggc ggggcetggt ggcegeggg 1980 egetgetggg tteteegagg egacetggee geeggeeget ceteegegeg etgtteegea 2040 cttgctgccc tcgcccggcc cggagcgccg ctgccatgcg gctggcgctg ctctqgqccc 2100 tggggctcct gggcgcgggc agccctctgc cttcctggcc gctcccaaat ataggtgagt 2160 cetecgeetg gagtgggteg gggggeggae tggggagggag gtgeaggaaa gteggaagge 2220 attagggtaa tggggccgga cggagaccct gggagagccc agccagagcg cggcccgcc 2280 tggtccgctg tcctgggcct agggcccgqt gacttggcga tggggtgaaa agagaaggag 2340 gggggatgcc ggcgcccct gcctcctgcc tggtcatcct ctgcgcggtc cctgcggaca 2400 ctttcaggct caggtaccag gtaccgaggg gcctgtccag cgccacttca agatcgtgat 2460 gagagggtcg ctqctcccca qqactqqcat cttcqctqct ctqqqqccta qctaaccqtt 2520 ccacccggtg ccagggcgct gagcgggcat ggcttgtagg gtttagtgaa gaggattctc 2580 tetageetet attecaggee tggggeegee aggeacteet caccetggtg etgttgeeae 2640 cagtgcctgg ccgagcggga ggggcccgag atgagccagg agaagggaga attggccagg 2700 aaagaggctg ggacaccaac tecteettgg aacttteact tecegetget gtettgggee 2760 gggaccgaga gggcaggcgc gggtggagtg tccggaggag agagggccat tgtgtttgg 2820 gggggtgggg ggtgctcgag gaggaagcag aggctgtagg cagcgggtgt gcctgactgg 2880 gcatgagggt gtttaggggg gtgggggtgt ttgcactgct cacccagaaa tgggcgttcc 2940 tggcatctcc gatgtgagcg aaggggaggg tgagcgggca cccggccaca aggcttagct 3000 cagtctcgag agggggcgtt cctgaagtgg ggggagagtg attgggaggg agtgggaacc 3060 geggagggte etgtgagaac etgggattgg eeggaagggg acaaggaggg ceacaggetg 3120 cgcaagccga aagtctttct tqqqqacttq tqaatqqqtt qqtqqqa aaqccataaa 3180 ttagagagac acceteteet teeagtatte ttetttaagt eteageatge aatgtggaag 3240 cccctcaggt acctaagggt cttgatgggc tgggagctgg tggatctgag ggcacctgtc 3300 acceccagee etgetgtega tteeeteagt actgtettgg ggtgteetgg gaeetgeagg 3360 tggcactgag gagcagcagg cagagtcaga gaaggccccg agggagccct tggagcccca 3420 ggtccttcag gacgatctcc caattagcct caaaaaggtg cttcaggtga gctctcactc 3480 ccctctaata aataaacgaa tccacacacg ccccqqtata gccaqqtqtc tcaaaqccaa 3540 agettggetg aggagetggt gggtagaget caetgtagtg ggtetatece aggeecaget 3600 geeteteeca ceacaceeca geacetgget teaettatet eceteteect etgeacacae 3660 gtgtatctgt ctgcctcagc cccacccaac ccatccatct ccactgggga aattgtgaag 3720 ccaaacttgc tttcttcatc tcatgttgtc ggttttctca gtggggggat ttggaaagag 3780 tcaggacett accaaacece ecceeccae eccattetaa agetgagtea gaggaaggge 3840 tggggcttgt gctgggtcct acacggtgct tcctctctgg gcaggaagcc gagaaggggt 3900 ggctcagata cetteettga ceteegcaca caaceeecca gaacaatget ceaggecagg 3960 cagggtttcc tggcccctcc cctgggatcc ccccaccagt gatctaattg ctggtgctct 4020 tetgtgggcc tgaggttttc tggttagaga ggctgggagt tgtggacagg tctagggagg 4080 tgacctgccc tctggtgccc acagaccagt ctgcctgagc ccctgaggat caagttggag 4140 ctggacggtg acagtcatat cctggagctg ctacagaata ggtaatagtg atggtggcaa 4200 taacagtgac cacatggcca acaacttgta tagcatttat tatgtgccag gtactaagtg 4260 cttgtgctca tttaatcctc ataacagccc tataagggat atactatcat gtattattgt 4320 cctcacttta tacatgagga agtcaaggca cagagagatt aaataacttg ccccaggtca 4380 cacagctagt atgtggtgaa aaccagattg gaattcaaat aaactaacag agtcagtggc 4440 ccaaccagta tactttgctg ccccggggtc aggagtggaa aagttggctg cgggggttgc 4500 etggteecca geeceacaac cacetteaag cetetgettg teaatgeace gaeectggga 4560 agtggcttta gcactgcctt ctttttcttc acttcacagg ggagttggtc ccatgtccgc 4620

congacont ggggtongs interested accounting generates the the 4680 tttcttcccc tccgctttcg tccttttgcc tcccccgtgc cgttgcgcgt tccttcttcc 4740 cegtteeete teeeetettt tgtteeetee egttetttte teeeeegegt tettteetee 4800 teettttegg teegeeeteg cetteetee tteecettet geeettegee nttteteet 4860 etegttette eteggtgteg egtegteeeg geteggeett teeeegette etecegeteg 4920 cogttttttt coccogctg tottocogtg troccottcg cttctcctct tccctttcqt 4980 teggtegttt tetegtteea tteeegeete eeegttteeg tteeacteet tetteeteet 5040 ttcccgctcc ccgtttctcc cgaccccaac aacaaataaa nnnnnnnnn nnnnnnnnn 5100 nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnntcagg aggccgagtg gaagaatcgc 5520 ttgagcccag gtaggcagag gtttcagtgg gccgagatcg agccactaca caccagcctg 5580 gctgagacag attaggtggt ttgcccgagg ccctacaact aataaatggc ctatccattt 5700 attagttgta tttggctctt catctgtctt atgatcccat ttgcagagag ctctcacttg 5760 gttatagata atacatagtt accaatqatg aagcaatata aacccaattt cctaatttgt 5820 aaaatgaaga taataaaact acttgctgca tagagttgct qqqaagatta aataagtcca 5880 tatagatgta aagtgcttaa aactatgcca gacctatggt aagtgacaag agttgttatt 5940 gggattttta aaattattat tattattatt attattatt gagacagagt ctcgctctgt 6000 ctcccaggct ggagtgcagt ggcgtgatct cggctcactg caagctccgc ctcccaqqtt 6060 cacgccattc tettgeetca geetceegag tagetgggae tacaggegee egecactaca 6120 cccggctaat gttttgtatt ttttagtaca gacagggttt caccgtgtta tccaggatgg 6180 tctcgatctc ctgacctcat gatccacccg ccttgtcctc ccaaagtgct gagattacag 6240 gcgtgagcca ccgcacccag ctaaattact gttttttaaa aatttgaaaa aaaccactga 6300 gtttggagcc agaaaagcag gggtctactc caaccttcat tatctacttc ctggtcctcc 6360 ttggcaagtt cctgggccct ctggccttca qtgqctcatc tqtaaaatqq qctcttcacc 6420 ctcctatttg acccacagag taggagtggc tgcctcttqg tcagcccggc acagctgctg 6480 gctgcgagcg gcaggtttgc ctgataattc ttcttgtcca tagtagaggc gggatgtggt 6540 aacagagacc aagactgtgg agttggtgat tgtggctgat cactcggagg tgagcctgct 6600 ggcccctgca catcctcctc cccctgcact gccctgccgc ctttcatgtc acctctcttg 6660 gectacagge ecagaaatae egggaettee ageacetget aaaceqeaca etqqaaqtqq 6720 ccctcttgct ggacacagtg agtgctggac agggcaaccc ccaccccagg cccctgacca 6780 tggcaacccc tcttctgagc cccagctgtc tttcagttct tccggcccct gaatgtacga 6840 gtggcactag tgggcctgga ggcctggacc cagcgtgacc tggtggagat cagcccaaac 6900 ccagctgtca ccctcgaaaa cttcctccac tggcgcaggg cacatttgct gcctcgattg 6960 ccccatgaca gtgcccagct ggtgacgtaa gggccccaga ctcagccaga gaggccagtc 7020 ctgtcctggc caaattcaca ccccttcagc accctacctc agcccctgaa gctctgacca 7080 ccgtggcttc tggccctgaa ctttagcctc tctgtcccac agtggtactt cattctctgg 7140 gcctacggtg ggcatggcca ttcagaactc catctgttct cctgacttct caggaggtgt 7200 gaacatggtg agttatttcc aggtctcctc ctcattccca attcagttcc tcccaagtgt 7260 ggtggcattt atgcactgaa accccctat aaagttgccc aaccccaaag ctacaggtat 7320 agagggtgga ggtacgtgat gtggcctttg ctatcaggga gccctcqctt atgqccaqct 7380 agtcacagtg tacacagtca tcccctgtgc agtcttccca tttcttaqaq qaqqqtaqqa 7440 ggcagctaag gcccaaagaa cagaggtgat ctccctccag tgagggaggg ggacagagct 7500 gagctagaac ccaagtttct gccatccagg cctgggttct cctactttag aagcaattca 7560 ggagggaagc agtgcctgct gagtgcccac gaggtcagac gtggagggaa caggagcaga 7620 gagggtggtc tgggcattgt ggtggaggca ggctgggact ggacctacag tacccctccc 7680 caatgacagg accactccac cagcatcctg ggagtcgcct cctccatagc ccatgagttg 7740 ggccacagcc tgggcctgga ccatgatttg cctgggaata gctgcccctg tccaggtcca 7800 gccccagcca agacctgcat catggaggcc tccacagagt aagtagctgc aggatggaga 7860 gagggtgtgg ggcagggggc agggannnnn nnnnnnnnn nnnnnnnnn tgttagagtt 7920 accttecttg ceaecetece eagetteeta ecaggeetga actteageaa etgeageega 7980 cgggccctgg agaaagccct cctggatgga atgggcagct gcctcttcga acggctqcct 8040

agcctacccc ctatggctgc tttctgcgga aatatgtttg tggagccggg cgagcagtgt 8100 gactgtggct tcctggatgt gagccccttt cccaaagcct cgccccactc acttctgtac 8160 cctcaccctg gctcattagc cctatcccag cctcctgagc tcttgggttc tgaagggact 8220 ttccacccct ctcctacttg ccctgtctgt ggggacagca catgggttgt tgggctctag 8280 ccctcqcttg ctgtgtagct tctggtcttg qcctgtggga ggaggagaga ttggagggag 8340 gctcacaggc cccacctgct ctgatgcccg gcccccgtgc tcctgcccac aggactgcgt 8400 cgatccctgc tgtgattctt tgacctgcca gctgaggcca ggtgcacagt gtgcatctga 8460 cggaccctgt tgtcaaaatt gccaggtggg tagagactag actggccacc cggagctcac 8520 ctgccggggc caaggtggaa agggtcattc tgacccccgg ctggatttgc tcagtgccca 8580 cactgatgct catccaccct ccacagctgc gcccgtctgg ctggcagtgt cgtcctacca 8640 qaqqqqattq tqacttqcct qaattctqcc caqqaqacaq ctcccaqtqt ccccctqatq 8700 teagectagg ggatggegag eeetgegetg gegggeaage tgtgtgeatg eaegggegtt 8760 gtgcctccta tgcccagcag tgccagtcac tttggggacc tggagcccag cccgctgcgc 8820 cactttgcct ccagaccgct aatactcggg gaaatgcttt tgggagctgt gggcgcaacc 8880 ccagtggcag ttatgtgtcc tgcaccccta ggtaagtgag gaaacctggc tcctcctttg 8940 ggtttctgag agccttggcc ctgctcctac taactctgtg tgcccttccc cctcnnnnnn 9000 tactcacact tttqccttca nacagctaat actcqqqqaa atqcttttqq qaqctqtqqq 9120 cqcaacccca qtqqcaqtta tqtqtcctqc acccctaqqt aaqtqaqqaa acctqqctcc 9180 tcctttgggt ttctgagagc cttggccctg ctcctactaa ctctgtgtgc ccttcccct 9240 ccccacagag atgccatttg tgggcagctc cagtgccaga caggtaggac ccagcctctg 9300 ctgggctcca tccgggatct actctgggag acaatagatg tgaatgggac tgagctgaac 9360 tgcagctggg tgcacctgga cctgggcagt gatgtggccc agccctcct gactctgcct 9420 ggcacagcct gtggccctgg cctggtgagc agcctgggtg ggcaagacca ggtgtgagaa 9480 gggacatttg gaccacaatg aacagagcc agacttcacc attcaccaat gtcaaaggca 9540 gggactccaa gggaagtcag tttcttactt cagatggagc aaagtcctat caactcacta 9600 tgccttggtt tccccatctg taaacgcagg gtatggcctc aaccttattg gcctcccagt 9660 cccattaaag ctttgtggga atctgatcca ggctcttctc tccctgggtc aggtgtgtat 9720 agaccatcga tgccagcgtg tggatctcct gggggcacag gaatgtcgaa gcaaatgcca 9780 tggacatggg gtgagctggg atgggggaag tggaagggga gcagagagcc tctagagagg 9840 aaaaggatac tgggctttgg aaatagacat atctgggttt taatccttgc tctactactt 9900 cccagttgtg tgacctcggg caggttacta actttgctga gctcagtttc cccacctatc 9960 aaatggctat aataatagta tccccatcca gggtacatga gatgtgtatg caagcaagta 10020 gcacagtggg taactaatag tgcttttaaa aannnnnnn nnnnnnnnn nnnnnnnnn 10080 



```
gaggttccgc tgcctccaag ccggacttag ggcttcaaga ggcgggcgtg ccctctggag 14940
tecectacea tgaetgaagg egecagagae tggeggtgte ttaagaetee gggeaeegee 15000
acgcgctgtc aagcaacact ctgcggacct gccggcgtag ttgcagcggg ggcttgggga 15060
ggggctgggg gttggacggg attgaggaag gtccgcacag cctgtctctg ctcagttgca 15120
ataaacgtga catcttggga gcgttcccca gagtttgtct gcttctagaa cccgggtcgc 15180
tectgetgeg gtteeaggtt tggeegeeag aagaegetge egeeteagae gagggeggge 15240
tgtgtggggc gggagtacca gaaagggtcg gcgtgtgtcc ccgggatgct cgcagcttcc 15300
ctctgcccag actggggtgg ctttcggcgc aatctgtcaa gctgttggac ctgccgtccc 15360
cactetgace attggctggg aaaagtggat etggctgatg etcecagage ecaggageca 15420
gggcggagcg gggcggcgc tgctcccacg atcccaaggc cgcgcacctg cctcctcccc 15480
ctccgccgcc gccacttgag ggatcgggaa caaaggtgct ttgtacaggc cgcaaccacc 15540
tcattacttc gtcttaggga ctggggccgc gtgggccccc agcccggaac gaaggtgtgg 15600
agcggcaagg gacagacgcc aatcttaaag tgagcatcta gcgcgccacc taaggctctt 15660
tagggaaggt ggtcccagag ctgtqttgtc ccttccqctt gcactgtccc tagatgtqca 15720
aagaaaacgg ggcagtgcat gaaggtggtt ggacaggctt catggatcct cgcccgcgcc 15780
tcactttccc ctatctgggc aaaggttatg tacccttatt taaaatcttc caaacttcta 15840
ataaggcagt ctaccctgca ctaaagcaga cacgaaagag atgacctccc taaaaaatact 15900
gctgttggaa tacgtccttc cttcccqccc cctcqcaqtq cqqtqcaqcc tcaqtqqaaq 15960
ctttggcgaa cctggcgcgc gctgcggtgc acagagggtt aactggagtt ggcgctgggt 16020
ggagaggagg agacgcgctc ccattggcgg aaagttattc aggggcgggg tcagtgaatc 16080
teegtaceee acteceettt eegeaaette eetetteaet ttgtacettt eteteetega 16140
ctgtgaagcg ggccgggacc tgccaggcca gaccaaaccg gacctcgggg gcgatgcggc 16200
tgctgcccct gctgcggact gtcctatggg ccgcgtcctc ggctcccctc tgcgcggggg 16260
ctccagcetc cgccacgtag tetactggaa etccagtaac eccaggtage egggeegaac 16320
cgggcgagcg cacagccaag tctgcgcgct cccgggcttt gcgcgcgccc gccacccgct 16380
ctttgcgcgg cgccgcctga gcctggccgc gcgccggggc tcctttgttt gaqccggcgg 16440
gggaggggg agggggagg ggcgaggcgc gccctgggtc tccccacagc ccgcatgtgt 16500
tggggggcag gcagaagacc ccagccccaa gggttgtcta gggggtcttg gagcatggag 16560
ctgggggggc ctttgcccgc actccgggct ccgccccct cgctgctctc ctggcgatcc 16620
ccagcctccc gcaggctgga gctgtggctg acgaacttga gagcgaggga gggggcttta 16680
ctcttatgaa agagcgtggg ttactctcct gcccgctggg tctcacctct ggctctcact 16740
ctgtctcctg atctcatttg ctatctctgc tttcatctct gtctttattg gtccttctgt 16800
ttctttccag tqtcagccct qcccttctag ccgaatcacc tctgggcaag tctcqtgacc 16860
ttcctaacct catttatctc acctgtataa tgggctaata atacctagta ccctgggaag 16920
tctggcaggg taagtgaggt catgtatgtg aaagaggctc aggctgtaca gatataaact 16980
attatttett tetetetet gagetgeetg cetttgaace ttagtatatt ttactgttte 17040
catecceete eccaagtete ectgeetete etatteeta tetgtttte tttetgattt 17100
tctacttgag acaatctgtg actattcatt tcttcact
                                                                  17138
```

Gln Ala Glu Ser Glu Lys Ala Pro Arg Glu Pro Leu Glu Pro Gln Val 40 Leu Gln Asp Asp Leu Pro Ile Ser Leu Lys Lys Val Leu Gln Thr Ser 55 60 Leu Pro Glu Pro Leu Arg Ile Lys Leu Glu Leu Asp Gly Asp Ser His Ile Leu Glu Leu Leu Gln Asn Arg Glu Leu Val Pro Gly Arg Pro Thr Leu Val Trp Tyr Gln Pro Asp Gly Thr Arg Val Val Ser Glu Gly His 100 105 Thr Leu Glu Asn Cys Cys Tyr Gln Gly Arg Val Arg Gly Tyr Ala Gly 120 Ser Trp Val Ser Ile Cys Thr Cys Ser Gly Leu Arg Gly Leu Val Val 135 Leu Thr Pro Glu Arg Ser Tyr Thr Leu Glu Gln Gly Pro Gly Asp Leu 150 155 Gln Gly Pro Pro Ile Ile Ser Arg Ile Gln Asp Leu His Leu Pro Gly 165 170 His Thr Cys Ala Leu Ser Trp Arg Glu Ser Val His Thr Gln Thr Pro 185 Pro Glu His Pro Leu Gly Gln Arg His Ile Arg Arg Arg Asp Val 200 Val Thr Glu Thr Lys Thr Val Glu Leu Val Ile Val Ala Asp His Ser 215 220 Glu Ala Gln Lys Tyr Arg Asp Phe Gln His Leu Leu Asn Arg Thr Leu 230 235 Glu Val Ala Leu Leu Leu Asp Thr Phe Phe Arg Pro Leu Asn Val Arg 245 250 Val Ala Leu Val Gly Leu Glu Ala Trp Thr Gln Arg Asp Leu Val Glu 260 265 Ile Ser Pro Asn Pro Ala Val Thr Leu Glu Asn Phe Leu His Trp Arg 280 275 Arg Ala His Leu Leu Pro Arg Leu Pro His Asp Ser Ala Gln Leu Val 295 300 Thr Gly Thr Ser Phe Ser Gly Pro Thr Val Gly Met Ala Ile Gln Asn 310 315 Ser Ile Cys Ser Pro Asp Phe Ser Gly Gly Val Asn Met Asp His Ser 325 330 Thr Ser Ile Leu Gly Val Ala Ser Ser Ile Ala His Glu Leu Gly His 345 Ser Leu Gly Leu Asp His Asp Leu Pro Gly Asn Ser Cys Pro Cys Pro 360 Gly Pro Ala Pro Ala Lys Thr Cys Ile Met Glu Ala Ser Thr Asp Phe 375 380 Leu Pro Gly Leu Asn Phe Ser Asn Cys Ser Arg Arg Ala Leu Glu Lys 390 395 Ala Leu Leu Asp Gly Met Gly Ser Cys Leu Phe Glu Arg Leu Pro Ser 405 410 Leu Pro Pro Met Ala Ala Phe Cys Gly Asn Met Phe Val Glu Pro Gly 425 Glu Gln Cys Asp Cys Gly Phe Leu Asp Asp Cys Val Asp Pro Cys Cys 435 440 Asp Ser Leu Thr Cys Gln Leu Arg Pro Gly Ala Gln Cys Ala Ser Asp 455 460 Gly Pro Cys Cys Gln Asn Cys Gln Leu Arg Pro Ser Gly Trp Gln Cys 470 475

Arg Pro Thr Arg Gly Asp Cys Asp Leu Pro Glu Phe Cys Pro Gly Asp 485 490 Ser Ser Gln Cys Pro Pro Asp Val Ser Leu Gly Asp Gly Glu Pro Cys 500 505 Ala Gly Gly Gln Ala Val Cys Met His Gly Arg Cys Ala Ser Tyr Ala 520 Gln Gln Cys Gln Ser Leu Trp Gly Pro Gly Ala Gln Pro Ala Ala Pro 535 Leu Cys Leu Gln Thr Ala Asn Thr Arg Gly Asn Ala Phe Gly Ser Cys 550 555 Gly Arg Asn Pro Ser Gly Ser Tyr Val Ser Cys Thr Pro Arg Asp Ala 570 Ile Cys Gly Gln Leu Gln Cys Gln Thr Gly Arg Thr Gln Pro Leu Leu 580 585 Gly Ser Ile Arg Asp Leu Leu Trp Glu Thr Ile Asp Val Asn Gly Thr 600 Glu Leu Asn Cys Ser Trp Val His Leu Asp Leu Gly Ser Asp Val Ala 615 620 Gln Pro Leu Thr Leu Pro Gly Thr Ala Cys Gly Pro Gly Leu Val 630 635 Cys Ile Asp His Arg Cys Gln Arg Val Asp Leu Leu Gly Ala Gln Glu 650 Cys Arg Ser Lys Cys His Gly His Gly Val Cys Asp Ser Asn Arg His 660 665 Cys Tyr Cys Glu Glu Gly Trp Ala Pro Pro Asp Cys Thr Thr Gln Leu 680 Lys Ala Thr Ser Ser Leu Thr Thr Gly Leu Leu Ser Leu Leu Val 695 700 Leu Leu Val Leu Val Met Leu Gly Ala Ser Tyr Trp Tyr Arg Ala Arg 710 715 Leu Xaa Gln Arg Leu Cys Gln Leu Lys Gly Pro Thr Cys Gln Tyr Arg 725 730 Ala Ala Gln Ser Gly Pro Ser Glu Arg Pro Gly Pro Pro Gln Arg Ala 745 Leu Leu Ala Arg Gly Thr Lys Ser Gln Gly Pro Ala Lys Pro Pro Pro 760 Pro Arg Lys Pro Leu Pro Ala Asp Pro Gln Gly Arg Cys Pro Ser Gly 775 780 Asp Leu Pro Gly Pro Gly Pro Gly Ile Pro Pro Leu Val Val Pro Ser 790 795 Arg Pro Ala Pro Pro Pro Pro Thr Val Ser Ser Leu Tyr Leu 805